

UC Irvine

UC Irvine Previously Published Works

Title

Consistency of spatial patterns of the daily precipitation field in the western United States and its application to precipitation disaggregation

Permalink

<https://escholarship.org/uc/item/8z28s1jx>

Journal

Geophysical Research Letters, 38(4)

ISSN

0094-8276

Authors

Chu, W
Gao, X
Phillips, TJ
et al.

Publication Date

2011-02-01

DOI

10.1029/2010GL046473

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Consistency of spatial patterns of the daily precipitation field in the western United States and its application to precipitation disaggregation

Wei Chu,¹ Xiaogang Gao,¹ Thomas J. Phillips,² and Soroosh Sorooshian¹

Received 12 December 2010; revised 13 January 2011; accepted 24 January 2011; published 23 February 2011.

[1] We investigate spatial patterns of daily precipitation field in the western United States, in order to improve assessment and disaggregation of climate model simulation. Empirical Orthogonal Function (EOF) analysis reveals that the spatial pattern of daily precipitation has not changed saliently in the region over the period 1948–2008. Results show that, even at very fine spatial ($.25^\circ \times .25^\circ$) and temporal (daily) resolutions, a small number (~ 15) of leading EOFs can explain about 90% of the total variance of the time-series of the entire precipitation field, having more than 6,000 grid cells. Moreover, the identified leading EOFs demonstrate consistency over time and across different spatial resolutions. Utilizing this consistency, an empirical method of disaggregating the precipitation output of climate models in this region is introduced. Illustrative results exhibit the feasibility and potency of this method. The advantages and limitations of this method are discussed as well. **Citation:** Chu, W., X. Gao, T. J. Phillips, and S. Sorooshian (2011), Consistency of spatial patterns of the daily precipitation field in the western United States and its application to precipitation disaggregation, *Geophys. Res. Lett.*, 38, L04403, doi:10.1029/2010GL046473.

1. Introduction

[2] Precipitation, especially over mountain ranges, is a crucial source of freshwater in the western United States. Furthermore, in this region, the spatial distribution of total precipitation deposited does not match well with the distribution of water demands from metropolitan, agricultural, and industrial areas. Massive infrastructures, such as the Federal Central Valley project and California State Water Project, have been built to remediate the spatial imbalances between water supply and demand. Therefore, understanding precipitation spatial distribution in this region, especially at high-resolutions, is critical to ensuring a reliable water supply in the face of increased climate variability.

[3] Numerous studies of the precipitation patterns in the western U.S. have been reported [Cayan *et al.*, 1998; Dettinger *et al.*, 1998], which identified the north-south precipitation patterns at interannual to interdecadal time scales. These studies focused on investigating the physical mechanisms responsible for the observed precipitation pat-

terns, linking the modes of precipitation patterns to climatic forcings (such as sea level pressure and sea surface temperature) and atmospheric circulations at both regional and global scales. The analysis presented in these studies was conducted on precipitation data of low spatial and temporal resolution. Instead of explaining the physical processes underlying the precipitation patterns, the objective of the current study is to investigate the properties of these patterns resolved at high spatial and temporal resolutions. We thus conduct our analysis on a 0.25-degree gridded field of daily precipitation data over the region bounded by 110° – 130° W and 30° – 50° N.

[4] Empirical Orthogonal Function (EOF) analysis [see, e.g., Dunteman, 1989] is employed in this study to extract the dominant spatial modes in the precipitation field. EOF analysis has been used extensively in atmospheric research to identify dominant orthogonal modes of the variance of meteorological fields [Peltier and Tushingham, 1989; Wagner *et al.*, 1990; Mann *et al.*, 1998a, 1998b; Deser *et al.*, 1999; Gong and Wang, 1999; Chen *et al.*, 2002; Häkkinen and Rhines, 2004]. Many studies have confirmed that modes represented by dominant EOFs can be related to underlying physical mechanisms [Semazzi *et al.*, 1988; Kawamura, 1994; Finnigan and Shaw, 2000]. However, in this study, we treat EOF analysis as a tool to extract the dominant signals, which are indicated by the variance explained by the leading EOFs, in the timeseries of the precipitation field. The remaining EOFs represent relatively trivial contributions to the total variance resulting from several confounding sources such as measurement or processing errors, fine-scale noise, localized precipitation features, etc.

[5] We initiatively apply EOF analysis to the precipitation fields at very high temporal (daily) and spatial ($.25^\circ$) resolution in order to study the spatial patterns resolved at such fine resolutions. The results reveal that a small number (~ 15) of leading EOFs can explain about 90% of the total variance of the daily precipitation field over the western U.S. Moreover, the leading EOFs demonstrate consistency over time and across spatial resolutions. Encouraged by this result, we propose a novel EOF-based method for directly disaggregating coarse-grid, model-simulated precipitation.

[6] Recently, Maraun *et al.* [2010] provided a comprehensive review of the state of the art of precipitation downscaling. However, the disaggregating method in this study does not fall into any of the categories defined therein. In general, standard statistical downscaling methods formulate the links between precipitation and predictor variables, such as geopotential height, sea level pressure, geostrophic velocity, and wind speed. In contrast, our method directly disaggregates low-resolution precipitation based on the

¹Department of Civil and Environmental Engineering, University of California, Irvine, California, USA.

²Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, Livermore, California, USA.

Table 1. Subsets of Precipitation Data Used in the Experiments

Dataset	Source	Resolution	Period	Role
A	Retropective analysis	$1^\circ \times 1^\circ$	1998–2007	Historical observation at the typical GCM's spatial resolution
B	Real-time analysis	$.25^\circ \times .25^\circ$	1998–2007	Historical observation at the target spatial resolution
C	Retropective analysis	$1^\circ \times 1^\circ$	2008	Model output to be disaggregated
D	Real-time analysis	$.25^\circ \times .25^\circ$	2008	Ground truth for validation of the disaggregation results
E	Retropective analysis	$1^\circ \times 1^\circ$	1948–1997	Observation to substantiate the long-term temporal persistency of EOFs

consistency of spatial patterns, which is very similar to the conceptual basis of the singular value decomposition (SVD) method introduced by *Widmann et al.* [2003]. In their study, they disaggregated 150×200 km wintertime precipitation (1958–94) of NCEP–NCAR reanalysis [*Kalnay et al.*, 1996] to a 50×50 km resolution over Oregon and Washington. Our study demonstrates the feasibility and effectiveness of this direct disaggregation procedure at high spatial and temporal resolutions.

[7] Section 2 briefly describes the data and methodology. Results are presented in section 3. Finally, we conclude with discussion of the results in section 4.

2. Data and Methods

[8] Historical precipitation over this region is extracted from US_Mexico daily precipitation analysis (retrospective) and U.S. daily precipitation analysis (real-time) [*Higgins et al.*, 2000] obtained from the Climate Prediction Center (CPC) at the National Weather Service (<ftp.cpc.ncep.noaa.gov/precip/>). The retrospective analysis spans the period 1948 to 2008 with spatial resolution of $1^\circ \times 1^\circ$, and the real-time analysis spans the period 1996 to the present at a resolution of $0.25^\circ \times 0.25^\circ$. The datasets are derived from two sources: the River Forecast Centers (RFC) that include data from about 6000 gauge stations per day, and the Climate Anomaly Data Base based on several hundred gauge stations per day. Prior to March 4, 1998, only the RFC data, with about 3000–6000 gauge stations per day, are available. Consequently, for the 1996–1998 period, the total daily precipitation of real-time analysis does not agree well with that of the retrospective analysis over the region. Table 1 lists

five subsets of data from the retrospective and real-time analysis for our experiment. As in the case of *Widmann et al.* [2003], we use observational data to represent climate model simulation output in order to exclude the errors that are introduced by a model, thus providing an accurate estimate of the uncertainty associated with the disaggregation method itself.

[9] Meteorological variables are usually represented as time-varying fields of geographical grids, where the value at each grid point is viewed as a variable. For instance, 6561 (81×81) grids are needed to represent the precipitation field at a given time over the entire study region at $.25^\circ$ resolution. Consequently, the timeseries of this precipitation field is equivalent to a timeseries of 6561 variables. However, owing to the correlations among neighboring grid values, EOF analysis can reduce the number of variables required to represent and characterize the precipitation field. EOF analysis is a multivariate statistical analysis tool that transforms a given dataset to a new orthogonal, independent coordinate system so that the first coordinate (i.e. the first EOF) has the largest projection of the dataset's variance, the second coordinate the second largest projection, and so on.

[10] Utilizing the leading EOFs' consistency over time and spatial resolution, we have developed a new method for disaggregating model simulations of precipitation in the western U.S. As an illustrative exercise, the method is implemented as follows:

[11] 1. Conduct EOF analysis on precipitation observations at both low-resolution, dataset A, and high-resolution, dataset B. The leading 15 EOFs in each analysis are then used to form the matrixes V_l and V_h that each have 15 columns corresponding to the EOFs.

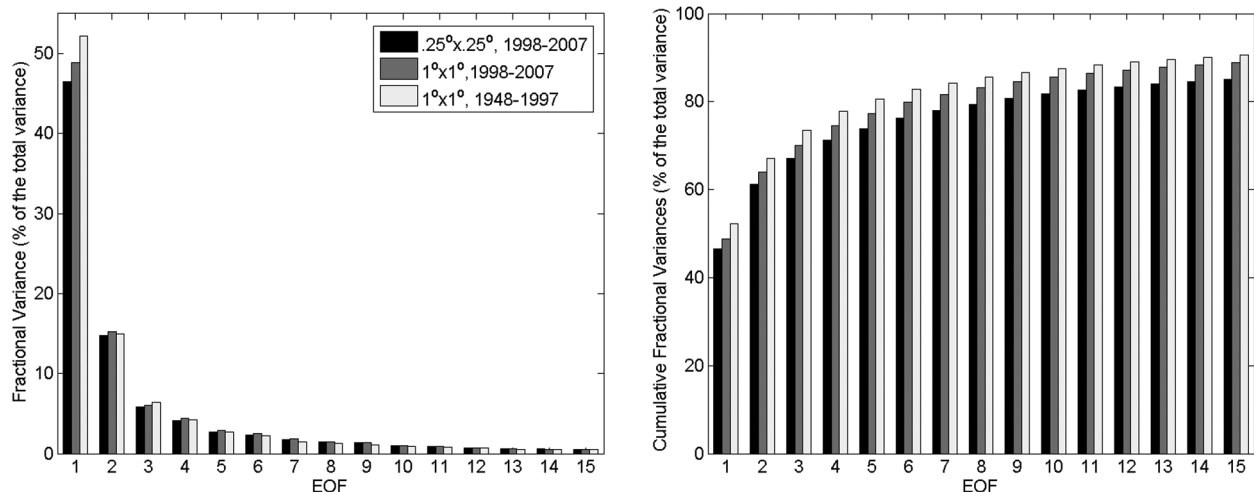


Figure 1. (left) Fractional variance and (right) cumulative fractional variance explained by the first 15 EOFs of precipitation fields of the $.25^\circ \times .25^\circ$ grid over 1998–2007, the $1^\circ \times 1^\circ$ grid over 1998–2007, and the $1^\circ \times 1^\circ$ grid over 1948–1997.

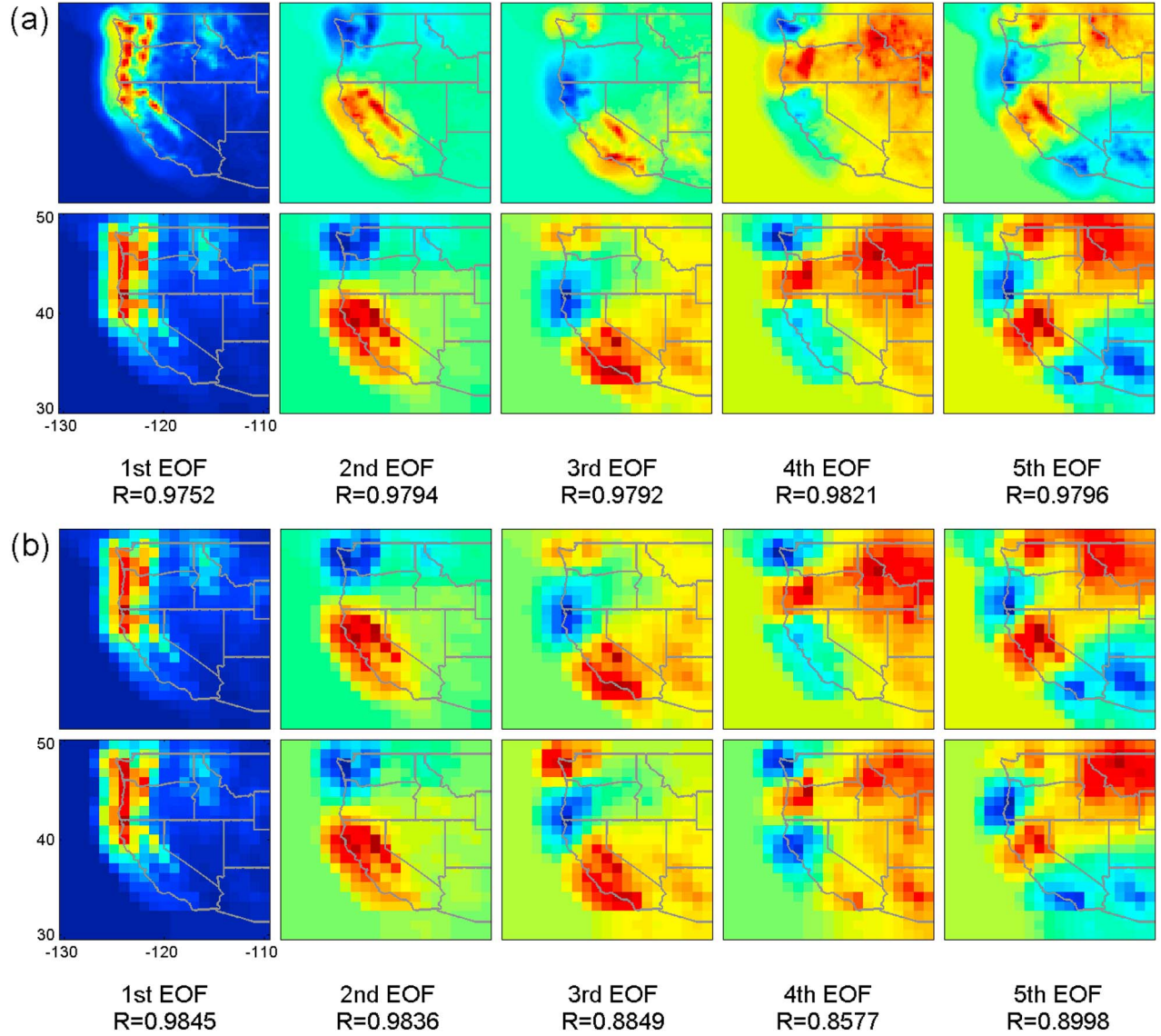


Figure 2. Comparison of the first five leading EOFs of (a) the precipitation field at different spatial resolutions: (top) $.25^\circ \times .25^\circ$ grid and (bottom) $1^\circ \times 1^\circ$ grid over the period 1998–2007; and (b) the precipitation on $1^\circ \times 1^\circ$ grid over two time periods: (top) 1998–2007 and (bottom) 1948–1997. R denotes the spatial correlation of corresponding EOFs.

[12] 2. Construct matrix A_l , where each row in A_l represents the timeseries of a grid-point value in dataset C. Therefore A_l has the size of $441(\text{number of grid points in the study region}) \times 366(\text{number of days in year 2008})$.

[13] 3. Given the EOFs' temporal consistency, assume that in a single year (e.g. 2008), the dominant EOFs of the previous 10 years (1998–2007) still dominate. Then, we can project A_l onto the low-resolution EOFs to obtain the principal component (PC) timeseries:

$$P_l = A_l V_l. \quad (1)$$

[14] 4. Since the EOFs are consistent across spatial resolutions, the PCs of high-resolution EOFs are strongly correlated to the PCs of low-resolution EOFs, as shown by the results described in Section 3 so that:

$$P_h \propto P_l. \quad (2)$$

[15] 5. Then, the high-resolution precipitation timeseries can be retrieved by projecting the PCs of high-resolution EOFs back onto the precipitation measurement space:

$$A_h = P_h V_h^T. \quad (3)$$

[16] 6. Now, A_h has the size of $6561(\text{number of high-resolution grid points in the study region}) \times 366$. Set negative values in A_h to zero, and normalize each column of A_h to make the total daily precipitation consistent (i.e. the sums of the corresponding columns in A_h and A_l are equal).

3. Results

[17] EOF analysis is conducted on datasets A, B, and E of Table 1. Results show that despite the differences in spatial resolution and time period, three datasets yield very similar EOF spectra (Figure 1, left). The corresponding leading

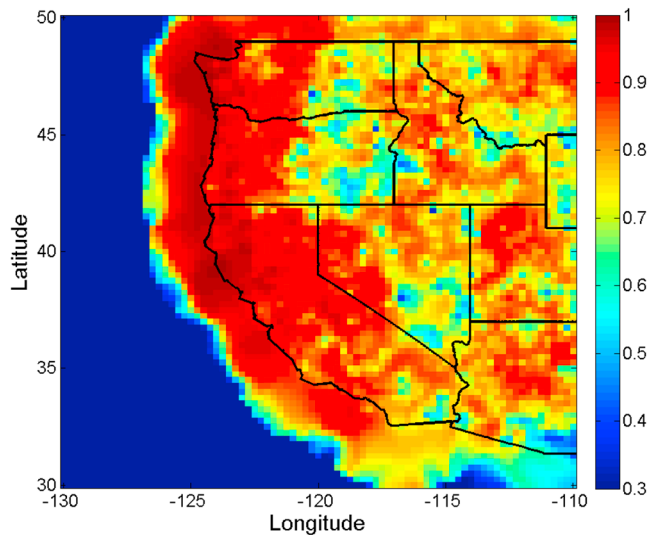


Figure 3. Correlations of the $.25^\circ \times .25^\circ$ daily precipitation disaggregated by the EOF method and the real-time analysis for the year 2008.

EOFs of the three datasets explain very similar amounts of fractional variance. The leading EOFs are highly dominant, with the first 15 explaining about 90% of the total variance (Figure 1, right). This result implies that the timeseries of 15 EOFs retain the major part of the information (variance) conveyed by the timeseries of the precipitation field.

[18] The consistency of EOF patterns across different spatial resolutions is revealed by comparing results from datasets A and B, which span the same time period, but are represented at, respectively, $1^\circ \times 1^\circ$ and $0.25^\circ \times 0.25^\circ$ spatial resolutions. The first five EOFs of both datasets are plotted in Figure 2a. It is evident that the corresponding EOFs resemble each other very well. Furthermore, quantitative analysis is used to substantiate the resemblance perceived by eye. Pixel-wise correlation is a rigorous measure

of similarity between two digital images. The correlations between the corresponding EOFs are computed by a two-step procedure: 1) for the $.25^\circ \times .25^\circ$ EOFs, the arithmetic mean of each 4×4 pixel group is calculated to form an image of $1^\circ \times 1^\circ$; and 2) the pixel-wise correlation is calculated between this new image and the corresponding EOF of dataset A. The leading EOFs at two resolutions are found to exhibit very high correlation coefficients (Figure 2a). Furthermore, the PC timeseries corresponding to the leading EOFs are also highly correlated, where the coefficients are higher than 0.98 for the first seven PCs and the average coefficient of the 15 PCs is greater than 0.90. To demonstrate the consistency of EOF spatial patterns over time, leading EOFs from datasets A and E are compared. Both datasets have the same $1^\circ \times 1^\circ$ spatial resolution, but different time coverage. The resemblance of the corresponding EOFs is manifested by Figure 2b and by the high pixel-wise correlation coefficients. These results indicate that there is no detectable change in the spatial pattern of precipitation over the period 1948–2008.

[19] Correlations of disaggregated (from dataset C) and observed (dataset D) daily precipitation of 2008 on the 0.25° grid are shown in Figure 3. Compared with the results of the analysis using the SVD method by *Widmann et al.* [2003], our analysis yield much higher correlation coefficients over the entire region. For all the pixels, the correlation coefficient is 0.89, and the root mean square error (RMSE) is 1.6285 mm/day. Since the method conserves the total amount of precipitation (step 6, above), the bias between the disaggregated and observed values is exactly the difference between the total volumes of datasets C and D, which is not related to the skill of the method.

[20] The fractional variance that is not explained by the first 15 EOFs can be treated as a measure of uncertainty. Let B^* denote the precipitation timeseries retrieved by the first 15 EOFs and PCs of dataset B, according to equation (3). Then $U = B - B^*$, represents the uncertainty field. Since there are 6561 pixels over the region, U contains 6561 timeseries. Next, we calculate the standard deviation (SD) of

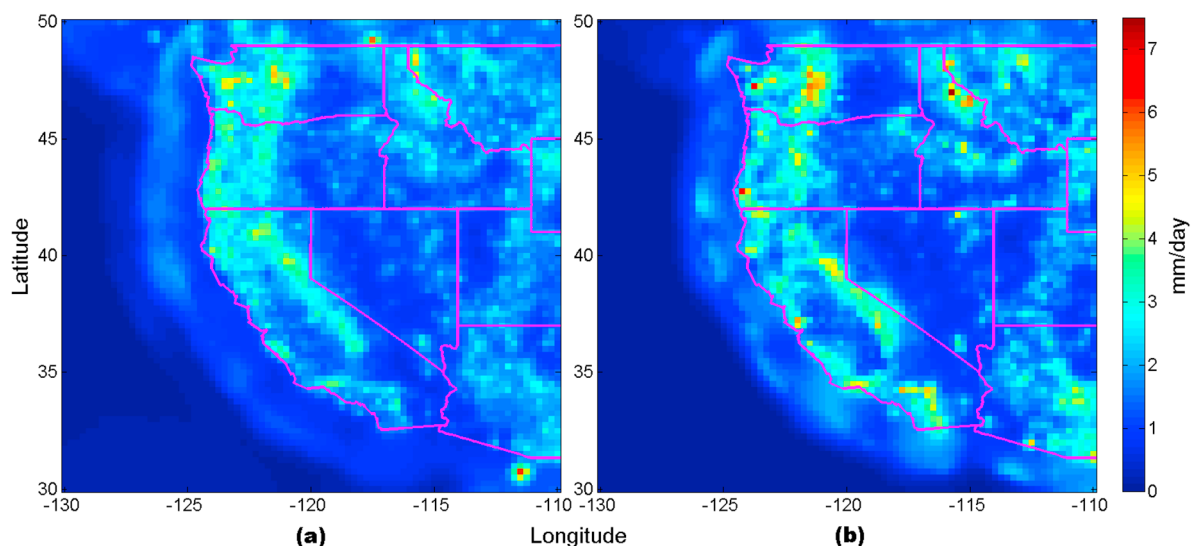


Figure 4. Maps of standard-deviation measures of (a) the uncertainty of the disaggregated precipitation values estimated from historical data, and (b) the residuals of the disaggregated result against the observation for 2008.

each individual timeseries to quantify the uncertainty at the corresponding grid point, yielding a SD map over the region (Figure 4a). The SD map provides uncertainty estimation for the disaggregated precipitation values at each grid point. To validate this estimation, the SD map of the residual R in 2008, which is defined as $R = \text{value disaggregated from dataset C} - \text{observed precipitation (dataset D)}$, is also constructed in the same way (Figure 4b). The two SD maps agree very well with each other. The mean SD values are 1.30 mm/day for U and 1.27 mm/day for R. The spatial correlation coefficient of the two maps is 0.88, and the RMS difference is 0.22 mm/day.

4. Discussion

[21] In this study, spatial patterns of precipitation resolved at fine temporal and spatial resolutions over the western U.S. are retrieved through EOF analysis. It is demonstrated that the variance of the entire high-resolution precipitation field can be approximated by a limited number of leading EOFs. Even though it may not be easy to explicitly link the EOFs with individual physical processes, these mathematical structures can still be productively used for practical applications.

[22] The persistency of the EOF patterns over the past half century suggests that there has not been significant change in the spatial pattern of daily precipitation over the region. Along with the consistency of leading EOFs across different spatial resolutions, the persistent spatial pattern allows maximum use of available observations for disaggregating model simulations of precipitation. Since models maintain precipitation fields' spatial and temporal coherence, our approach does not need to include procedures for restoring coherence, such as the Schaake Shuffle [Clark *et al.*, 2004], that are required by many statistical downscaling methods. The approach also separates the dominant signals from the trivial, allowing these two components with distinct characteristics to be treated separately and properly.

[23] On the other hand, the EOF disaggregation method also has its limitations. In order to take advantage of the consistency of spatial patterns, the method requires long-term historical precipitation data. The historical data should include the precipitation observations at both the resolution of the model and of the target, such as Table 1 datasets A and B. In addition, the method assumes that a model-generated precipitation field is reasonably accurate, so that the coarse-grid, simulated precipitation can be confidently redistributed on a finer mesh. In practice, a validation and bias-correction of model-generated precipitation is required before carrying out the steps of the procedure outlined in Section 2.

[24] In fact, validation and bias-correction are other potential applications of the EOF patterns revealed in this study, as demonstrated by Nieto and Rodríguez-Puebla [2006] and Biau *et al.* [1999]. Our current focus is on developing new metrics and procedures to evaluate climate models based on their ability to simulate the consistent precipitation patterns. It is expected the results will guide us in selecting proper models to generate reliable ensemble projection of precipitation in the region. The uncertainties associated with this ensemble projection will also be quantitatively delineated.

[25] **Acknowledgments.** We are very grateful to two anonymous reviewers for their valuable and constructive comments. This research was supported by UCOP program of University of California (grant 09-LR-09-116849-SORS), CPPA program of NOAA (grants NA08OAR4310876 and NA05OAR4310062), and ROSES program of NASA (grant NNX09AO67G).
[26] The Editor thanks one anonymous reviewer.

References

- Biau, G., E. Zorita, H. von Storch, and H. Wackernagel (1999), Estimation of precipitation by kriging in the EOF space of the sea level pressure field, *J. Clim.*, *12*, 1070–1085, doi:10.1175/1520-0442(1999)012<1070:EOPBK1>2.0.CO;2.
- Cayan, D. R., M. D. Dettinger, H. F. Diaz, and N. E. Graham (1998), Decadal variability of precipitation over western North America, *J. Clim.*, *11*, 3148–3166, doi:10.1175/1520-0442(1998)011<3148:DVOPW>2.0.CO;2.
- Chen, J. Y., B. E. Carlson, and A. D. D. Genio (2002), Evidence for strengthening of the tropical general circulation in the 1990s, *Science*, *295*, 838–841, doi:10.1126/science.1065835.
- Clark, M. P., S. Gangopadhyay, L. Hay, B. Rajagopalan, and R. Wilby (2004), The Schaake Shuffle: A method for reconstructing space–time variability in forecasted precipitation and temperature fields, *J. Hydro-meteorol.*, *5*, 243–262, doi:10.1175/1525-7541(2004)005<0243:TSSAMF>2.0.CO;2.
- Deser, C., M. A. Alexander, and M. S. Timlin (1999), Evidence for a wind-driven intensification of the Kuroshio Current Extension from the 1970s to the 1980s, *J. Clim.*, *12*, 1697–1706, doi:10.1175/1520-0442(1999)012<1697:EFAWDI>2.0.CO;2.
- Dettinger, M. D., D. R. Cayan, H. F. Diaz, and D. M. Meko (1998), North–south precipitation patterns in western North America on interannual–to–decadal timescales, *J. Clim.*, *11*, 3095–3111, doi:10.1175/1520-0442(1998)011<3095:NSPPIW>2.0.CO;2.
- Dunteman, G. H. (1989), *Principal Components Analysis*, 97 pp., Sage, Newbury Park, Calif.
- Finnigan, J. J., and R. H. Shaw (2000), A wind tunnel study of airflow in waving wheat: An EOF analysis of the structure of the large-eddy motion, *Boundary Layer Meteorol.*, *96*, 211–255, doi:10.1023/A:1002618621171.
- Gong, D., and S. Wang (1999), Definition of Antarctic Oscillation index, *Geophys. Res. Lett.*, *26*, 459–462, doi:10.1029/1999GL900003.
- Häkkinen, S., and P. B. Rhines (2004), Decline of subpolar North Atlantic circulation during the 1990s, *Science*, *304*, 555–559, doi:10.1126/science.1094917.
- Higgins, R. W., W. Shi, and E. Yarosh (2000), Improved United States precipitation quality control system and analysis, *NCEP/Climate Prediction Center Atlas*, vol. 7, Natl. Cent. for Environ. Predict., Washington, D. C. 40 pp.
- Kalnay, E., *et al.* (1996), The NCEP/NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*, 437–471, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.
- Kawamura, R. (1994), A rotated EOF analysis of global sea surface temperature variability with interannual and interdecadal scales, *J. Phys. Oceanogr.*, *24*, 707–715, doi:10.1175/1520-0485(1994)024<0707:AREAOG>2.0.CO;2.
- Mann, M. E., R. S. Bradley, and M. K. Hughes (1998a), Global-scale temperature patterns and climate forcing over the past six centuries, *Nature*, *392*, 779–787, doi:10.1038/33859.
- Mann, M. E., R. S. Bradley, M. K. Hughes, and P. D. Jones (1998b), Global temperature patterns, *Science*, *280*, 2027, doi:10.1126/science.280.5372.2027e.
- Maraun, D., *et al.* (2010), Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, *Rev. Geophys.*, *48*, RG3003, doi:10.1029/2009RG000314.
- Nieto, S., and C. Rodríguez-Puebla (2006), Comparison of precipitation from observed data and general circulation Mmodels over the Iberian Peninsula, *J. Clim.*, *19*, 4254–4275, doi:10.1175/JCLI3859.1.
- Peltier, W., and A. Tushingham (1989), Global sea level rise and the greenhouse effect: Might they be connected, *Science*, *244*, 806–810, doi:10.1126/science.244.4906.806.
- Semazzi, F. H. M., V. Mehta, and Y. C. Sud (1988), An investigation of the relationship between sub-Saharan rainfall and global sea surface temperatures, *Atmos. Ocean*, *26*, 118–138.
- Wagner, D., E. Ruprecht, and C. Simmer (1990), A combination of microwave observations from satellites and an EOF Analysis to retrieve vertical humidity profiles over the ocean, *J. Appl. Meteorol.*, *29*, 1142–1157, doi:10.1175/1520-0450(1990)029<1142:ACOMOF>2.0.CO;2.
- Widmann, M., C. S. Bretherton, and E. P. Salathé (2003), Statistical precipitation downscaling over the northwestern united states using numeri-

cally simulated precipitation as a predictor, *J. Clim.*, 16, 799–816,
doi:10.1175/1520-0442(2003)016<0799:SPDOTN>2.0.CO;2.

T. J. Phillips, Program for Climate Model Diagnosis and Intercomparison,
Lawrence Livermore National Laboratory, PO Box 808, Livermore, CA
94551, USA.

W. Chu, X. Gao, and S. Sorooshian, Department of Civil and
Environmental Engineering, University of California, Irvine, CA 92697,
USA.